SUFFICIENT CONDITIONS FOR A STATIONARY PROCESS TO BE A FUNCTION OF A FINITE MARKOV CHAIN[1]

By S. W. Dharmadhikari

University of California, Berkeley

0. Summary. Let $\{Y_n, n \geqq 1\}$ be a stationary process with a finite state-space J. We will use the definitions of a function and of a regular function of a finite Markov chain given in [1].

In Section 1 of this paper we define, for each state ϵ of J, a convex cone $\mathcal{C}(\pi_\epsilon)$. The main theorem (Section 2) asserts that if each $\mathcal{C}(\pi_\epsilon)$ is polyhedral, then $\{Y_n\}$ is a function of a finite Markov chain. The hypothesis that each $\mathcal{C}(\pi_\epsilon)$ is polyhedral is not quite necessary and some results are given in Section 3 under weaker assumptions. It is also shown that these weaker conditions are necessary for $\{Y_n\}$ to be a *regular* function of a finite Markov chain. The final section presents an example which shows that not every function of a finite Markov chain is a regular function of a Markov chain.

The results of Gilbert [3] are at the root of our investigation. Gilbert always assumed that the given stationary process $\{Y_n\}$ and the underlying Markov chain were irreducible and aperiodic. However, his results continue to hold even when these assumptions are dropped. In particular, the results of Section 1 of [3] depend only on the stationarity and the Markov character of the underlying chain. Theorem 2 of [3] also holds in a more general set-up (see Lemma 3.1 of [1]).

Our stationary process $\{Y_n\}$ need not be irreducible or aperiodic. Our sufficient conditions also do not necessarily yield such a chain.

1. The cones $\mathcal{C}(\alpha_\epsilon)$ and $\mathcal{C}(\pi_\epsilon)$. Let $\{Y_n\}$ be a stationary process with a finite state-space $J = \{0, 1, \cdots, D - 1\}$. ϵ, μ will denote states of J and s, t (with or without affixes) will denote finite sequences of states of J. If $s = \epsilon_1 \cdots \epsilon_n$, let $p(s) = P[(Y_1, \cdots, Y_n) = s]$. We will assume that each $p(\epsilon)$ is positive. It is convenient to introduce the empty sequence denoted by ϕ, with the conventions that $p(\phi) = 1$ and $p(s\phi) = P(\phi s) = p(s)$. For every ϵ define $n(\epsilon)$ to be the highest integer n such that we can find $2n$ sequences s_i, t_i, $(i = 1, \cdots, n)$ of states of J such that the matrix $\|p(s_i \epsilon t_j)\|$ is non-singular. The numbers $n(\epsilon)$ were introduced by Gilbert [3]. *We will assume that each $n(\epsilon)$ is finite.* The finiteness of $n(\epsilon)$ implies ([1], Section 2) that, for every ϵ, there are sequences $s_{\epsilon i}, t_{\epsilon i}$, $(i = 1, \cdots, n(\epsilon))$ such that

(a) the matrix $\|p(s_{\epsilon i} \epsilon t_{\epsilon j})\|$ is non-singular and

Received October 15, 1962.

[1] Prepared with the partial support of the National Science Foundation, Grant G-14648. Part of the author's Ph.D. thesis submitted to the University of California, Berkeley. Work done while on leave from the University of Poona.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

(b) for each s, there are unique constants $a_{\epsilon i}(s)$ such that, for every t

$$p(s\epsilon t) = \sum_{i=1}^{n(\epsilon)} a_{\epsilon i}(s)p(s_{\epsilon i}\epsilon t). \tag{1.1}$$

Let E^k denote the k-dimensional Euclidean space. It is then evident that (1.1) expresses $p(s\epsilon t)$ as an inner product of two vectors in $E^{n(\epsilon)}$. Define the vectors $\alpha_\epsilon(s)$ and $\pi_\epsilon(t)$ in $E^{n(\epsilon)}$ by $\alpha_\epsilon(s) = (a_{\epsilon 1}(s), \cdots, a_{\epsilon n(\epsilon)}(s))$, $\pi_\epsilon(t) = (p(s_{\epsilon 1}\epsilon t), \cdots, p(s_{\epsilon n(\epsilon)}\epsilon t))$. Then (1.1) can be rewritten as

$$p(s\epsilon t) = (\alpha_\epsilon(s), \pi_\epsilon(t)). \tag{1.2}$$

Let α_ϵ denote the set of vectors $\alpha_\epsilon(s)$ for all s and let π_ϵ denote the set of vectors $\pi_\epsilon(t)$ for all t. Let $\mathcal{C}(\pi_\epsilon)$, $\mathcal{C}(\alpha_\epsilon)$ respectively be the closed convex cones generated by the sets π_ϵ and α_ϵ. We will now study some properties of these cones.

LEMMA 1.1. *For every ϵ, both $\mathcal{C}(\alpha_\epsilon)$ and $\mathcal{C}(\pi_\epsilon)$ have dimension $n(\epsilon)$.*

PROOF.

(i) For $i, k = 1, \cdots, n(\epsilon)$, observe that $a_{\epsilon i}(s_{\epsilon k}) = \delta_{ik}$. Therefore $\alpha_\epsilon(s_{\epsilon k})$ is the kth co-ordinate vector in $E^{n(\epsilon)}$. Hence $\mathcal{C}(\alpha_\epsilon)$ contains the non-negative orthant in $E^{n(\epsilon)}$. This proves that $\mathcal{C}(\alpha_\epsilon)$ has dimension $n(\epsilon)$.

(ii) The non-singularity of the matrix $\|p(s_{\epsilon i}\epsilon t_{\epsilon j})\|$ means that the vectors $\pi_\epsilon(t_{\epsilon i})$, $(i = 1, \cdots, n(\epsilon))$, are linearly independent. This proves that $\mathcal{C}(\pi_\epsilon)$ has dimension $n(\epsilon)$ and completes the proof of the lemma.

Let $\mathcal{C}$ be a cone in E^k. Then $\mathcal{C}^+ = \{y \text{ in } E^k \mid (x, y) \geqq 0 \text{ for all } x \text{ in } \mathcal{C}\}$ is called the dual cone of $\mathcal{C}$. Since probabilities are non-negative, it follows from (1.2) that

$$\mathcal{C}(\alpha_\epsilon) \subset [\mathcal{C}(\pi_\epsilon)]^+. \tag{1.3}$$

LEMMA 1.2. *Let b be a non-zero vector of $[\mathcal{C}(\pi_\epsilon)]^+$. Then $(b, \pi_\epsilon(\phi)) > 0$.*

PROOF. Suppose $(b, \pi_\epsilon(\phi)) = 0$. Observing that $\pi_\epsilon(\phi) = \sum_\mu \pi_\epsilon(\mu)$ and $(b, \pi_\epsilon(\mu)) \geqq 0$, it follows that $(b, \pi_\epsilon(\mu)) = 0$ for all μ. By induction, $(b, \pi_\epsilon(t)) = 0$ for all t. That is, b is perpendicular to $\mathcal{C}(\pi_\epsilon)$. But $\mathcal{C}(\pi_\epsilon)$ has full dimension, (by the previous lemma). Hence b must be the zero vector. This proves the lemma.

Let $a_{\epsilon i, \mu j} = a_{\mu j}(s_{\epsilon i}\epsilon)$ and let $A_{\epsilon\mu}$ be the $n(\epsilon) \times n(\mu)$ matrix whose (ij)th element is $a_{\epsilon i, \mu j}$. It follows easily from (1.1) that

$$\pi_\epsilon'(\mu t) = A_{\epsilon\mu}\pi_\mu'(t). \tag{1.4}$$

(1.4) implies that, for an arbitrary vector b in $E^{n(\epsilon)}$ and for every t,

$$(b, \pi_\epsilon(\mu t)) = (bA_{\epsilon\mu}, \pi_\mu(t)). \tag{1.5}$$

LEMMA 1.3. *For every ϵ, μ and s, $\alpha_\epsilon(s)A_{\epsilon\mu} = \alpha_\mu(s\epsilon)$.*

PROOF. Using (1.5) and (1.2), we have, for every t, $(\alpha_\epsilon(s)A_{\epsilon\mu}, \pi_\mu(t)) = (\alpha_\epsilon(s), \pi_\epsilon(\mu t)) = p(s\epsilon\mu t) = (\alpha_\mu(s\epsilon), \pi_\mu(t))$.
But $\mathcal{C}(\pi_\mu)$ has full dimension. Hence the lemma.

2. The main theorem. This section is devoted to showing that if each $\mathcal{C}(\pi_\epsilon)$ is polyhedral then $\{Y_n\}$ is a function of a finite Markov chain. We will first construct a transition probability matrix M. Next we will construct a stationary initial distribution $\mathbf{m}$ which, together with M, yields a Markov chain of which $\{Y_n\}$ is a function.

LEMMA 2.1. *For each* ϵ, *let* $[\mathcal{C}(\pi_\epsilon)]^+$ *be polyhedral and be generated by* $N(\epsilon)$ *non-zero vectors* $\beta_{\epsilon j}$, $(j = 1, \cdots, N(\epsilon))$. *Define* $r_{\epsilon j}(t) = (\beta_{\epsilon j}, \pi_\epsilon(t))$, *for every* t. *Then*

(i) *the* β*'s can be chosen in such a way that for each* ϵ *and for* $j = 1, \cdots, N(\epsilon)$,

$$r_{\epsilon j}(\phi) = 1; \tag{2.1}$$

(ii) *for all* ϵ *and* μ *and for* $j = 1, \cdots, N(\epsilon)$, $k = 1, \cdots, N(\mu)$, *there exist non-negative constants* $m_{\epsilon j,\mu k}$ *such that for all* t,

$$r_{\epsilon j}(\mu t) = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} r_{\mu k}(t); \tag{2.2}$$

and if (2.1) *holds, then*

$$\sum_{\mu=0}^{D-1} \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} = 1. \tag{2.3}$$

PROOF. Since the β's are unique only up to non-negative multiplicative constants, (i) will be proved as soon as we have $(\beta_{\epsilon j}, \pi_\epsilon(\phi)) > 0$. But this follows from Lemma 1.2.

Using (1.5) with $b = \beta_{\epsilon j}$, we get for every t,

$$(\beta_{\epsilon j} A_{\epsilon\mu}, \pi_\mu(t)) = r_{\epsilon j}(\mu t) \geqq 0. \tag{2.4}$$

Thus $\beta_{\epsilon j} A_{\epsilon\mu}$ belongs to $[\mathcal{C}(\pi_\mu)]^+$. Hence there exist constants $m_{\epsilon j,\mu k} \geqq 0$ such that

$$\beta_{\epsilon j} A_{\epsilon\mu} = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} \beta_{\mu k}. \tag{2.5}$$

(2.2) now follows by taking inner product of (2.5) with $\pi_\mu(t)$ and using (2.4) on the left side. If (2.1) holds, then putting $t = \phi$ in (2.2), we get

$$r_{\epsilon j}(\mu) = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k}. \tag{2.6}$$

Therefore $\sum_{\mu=0}^{D-1} \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} = \sum_\mu r_{\epsilon j}(\mu) = r_{\epsilon j}(\phi) = 1$, which proves (2.3) and completes the proof of the lemma.

Let the conditions of the previous lemma hold. Let $M_{\epsilon\mu}$ be the $N(\epsilon) \times N(\mu)$ matrix whose (ij)th term is the quantity $m_{\epsilon j,\mu k}$ given by the lemma. Let $N = \sum N(\epsilon)$ and define M to be the $N \times N$ matrix

$$M = \begin{pmatrix} M_{00} & \cdots & M_{0,D-1} \\ & \cdots & \\ & \cdots & \\ M_{D-1,0} & \cdots & M_{D-1,D-1} \end{pmatrix} \tag{2.7}$$

The non-negativity of the m's together with (2.3) shows that M is a transition probability matrix. The symbol I will denote a set of N elements conveniently enumerated as $\{\epsilon j \mid j = 1, \cdots, N(\epsilon); \epsilon = 0, 1, \cdots, D - 1\}$. Finally, f will denote a function on I onto J defined by $f(\epsilon j) = \epsilon$.

Let $\{X_n\}$ be a Markov chain with state-space I and transition probability matrix M. For a sequence t of length n, the quantity $r_{\epsilon j}(t)$ defined in the statement of the previous lemma is going to be interpreted as $r_{\epsilon j}(t) = P[(f(X_2), \cdots, f(X_{n+1})) = t \mid X_1 = \epsilon j]$. The Markov character of $\{X_n\}$ then requires that the Equation (2.2) be satisfied.

We will now proceed to find a stationary initial distribution for M. It will be convenient to write a vector x of N elements in the form $(x_0, \cdots, x_{D-1})$, where x_ϵ has the $N(\epsilon)$ elements $x_{\epsilon j}$, $(j = 1, \cdots, N(\epsilon))$.

(1.3) shows that $\alpha_\epsilon(\phi)$ belongs to $[\mathcal{C}(\pi_\epsilon)]^+$. Therefore, for every ϵ, we can find non-negative constants $m_{\epsilon j}^{(0)}$, $(j = 1, \cdots, N(\epsilon))$, such that

$$\alpha_\epsilon(\phi) = \sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{(0)} \beta_{\epsilon j}. \tag{2.8}$$

Let $\mathbf{m}^{(0)}$ be the vector of N elements defined by the $m_{\epsilon j}^{(0)}$. Taking inner product of (2.8) with $\pi_\epsilon(\phi)$, we get $p(\epsilon) = \sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{0}$. This shows that $\mathbf{m}^{(0)}$ defines a probability distribution on I. Define $\mathbf{m}^{(n+1)}$ for $n \geqq 0$ by induction as follows.

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} M. \tag{2.9}$$

By the standard theory of Markov chains it follows that the limit

$$\mathbf{m} = \lim_{n\to\infty}(1/(n+1)) \sum_{\nu=0}^{n} \mathbf{m}^{(\nu)} \tag{2.10}$$

exists and forms a stationary initial distribution for M.

LEMMA 2.2. *For $n \geqq 0$ and for every ϵ,*

$$\sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{(n)} \beta_{\epsilon j} = \alpha_\epsilon(\phi).$$

PROOF. (2.8) shows that the lemma holds for $n = 0$. Suppose it holds for $n = \nu$. Then, from (2.9), (2.5) and Lemma 1.3

$$\sum_{k=1}^{N(\mu)} m_{\mu k}^{(\nu+1)} \beta_{\mu k} = \sum_{k=1}^{N(\mu)} \beta_{\mu k} \sum_{\epsilon=0}^{D-1} \sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{(\nu)} m_{\epsilon j,\mu k} = \sum_{\epsilon=0}^{D-1} \sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{(\nu)} \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} \beta_{\mu k}$$
$$= \sum_{\epsilon=0}^{D-1} \sum_{j=1}^{N(\epsilon)} m_{\epsilon j}^{(\nu)} \beta_{\epsilon j} A_{\epsilon\mu} = \sum_{\epsilon=0}^{D-1} \alpha_\epsilon(\phi) A_{\epsilon\mu} = \sum_{\epsilon=0}^{D-1} \alpha_\mu(\epsilon) = \alpha_\mu(\phi).$$

The lemma thus follows by induction.

The next lemma brings together all the results we have obtained so far.

LEMMA 2.3. *Let M, $\mathbf{m}$ and f be as defined above. Let $\{X_n, n \geqq 1\}$ be a stationary Markov chain with state-space I, transition matrix M and initial distribution $\mathbf{m}$. Then $\{f(X_n), n \geqq 1\}$ has the same distribution as $\{Y_n, n \geqq 1\}$.*

PROOF. Lemma 2.2 shows that $\sum_{j=1}^{N(\epsilon)} m_{\epsilon j}\beta_{\epsilon j} = \alpha_\epsilon(\phi)$. Inner product with $\pi_\epsilon(t)$ yields

$$\sum_{j=1}^{N(\epsilon)} m_{\epsilon j} r_{\epsilon j}(t) = p(\epsilon t). \tag{2.11}$$

To prove the lemma it is now enough to show that for every n and for every t of length n

$$P[(f(X_2), \cdots, f(X_{n+1})) = t \mid X_1 = \epsilon j] = r_{\epsilon j}(t). \tag{2.12}$$

For, then (2.11) implies that $\{f(X_n)\}$ and $\{Y_n\}$ have the same distribution.

(2.6) shows that (2.12) holds for $n = 1$. Suppose it holds for $n = \nu$. If t has length $(\nu + 1)$, we can write $t = \mu t'$ for some μ and for some t' of length ν. Therefore from (2.2) and by stationarity,

$$\begin{aligned} P[(f(X_2), \cdots, f(X_{\nu+2})) = t \mid X_1 = \epsilon j] &= \sum_{k=1}^{N(\mu)} m_{\epsilon j, \mu k} P[(f(X_3), \cdots, f(X_{\nu+2})) = t' \mid X_2 = \mu k] \\ &= \sum_{k=1}^{N(\mu)} m_{\epsilon j \cdot \mu k} r_{\mu k}(t') = r_{\epsilon j}(\mu t') = r_{\epsilon j}(t). \end{aligned}$$

The lemma now follows by induction.

We are now in a position to state the main theorem of this paper.

THEOREM 2.1. *Let $\{Y_n, n \geqq 1\}$ be a stationary process with a finite state-space. Let $\sum n(\epsilon) < \infty$ and for each ϵ let $[\mathcal{C}(\pi_\epsilon)]^+$ be polyhedral and be generated by $N(\epsilon)$ non-zero vectors. Then $\{Y_n\}$ is a function of a stationary Markov chain with $\sum N(\epsilon)$ states.*

In view of the preceding lemma, this theorem does not need any proof. A special case of the theorem is of some interest and is given by the following corollary.

COROLLARY 2.1. *Let $\{Y_n\}$ be a stationary process with a finite state-space. Let $n(\epsilon) \leqq 2$, for every ϵ. Then $\{Y_n\}$ is a function of a stationary Markov chain with $\sum n(\epsilon)$ states—that is a regular function of a Markov chain.*

PROOF. We use the standard result that any convex cone in E^n is either the whole of E^n or is contained in a half-space. $[\mathcal{C}(\pi_\epsilon)]^+$ cannot be the whole of $E^{n(\epsilon)}$ because its dual, namely $\mathcal{C}(\pi_\epsilon)$, contains non-zero vectors.

(a) If $n(\epsilon) = 1$, then $[\mathcal{C}(\pi_\epsilon)]^+$ reduces to the non-negative real line and is clearly generated by 1 vector.

(b) Let $n(\epsilon) = 2$. Since $\mathcal{C}(\pi_\epsilon)$ is two-dimensional, $[\mathcal{C}(\pi_\epsilon)]^+$ must be a proper subset of a half-space. But every such cone in E^2 is generated by at most two vectors. $[\mathcal{C}(\pi_\epsilon)]^+$ cannot be generated by 1 vector because it has dimension 2.

Thus if $n(\epsilon) \leqq 2$, then $[\mathcal{C}(\pi_\epsilon)]^+$ is polyhedral and is generated by $n(\epsilon)$ vectors. The corollary now follows from the above theorem.

The result stated in this corollary has been reported by Fox [2].

3. A theorem under weaker assumptions. The theorem proved in the preceding section gives sufficient conditions for a stationary process to be a function of a finite Markov chain. It is therefore of some interest to examine whether the hypothesis that each $\mathcal{C}(\pi_\epsilon)$ is polyhedral can be weakened. The non-negativity of M was a consequence of the fact that $\beta_{\epsilon j}A_{\epsilon\mu}$ belonged to $[\mathcal{C}(\pi_\mu)]^+$. If we incorporate a similar property into our assumptions, we do not need to assume that $\mathcal{C}(\pi_\epsilon)$ is polyhedral. There is no difficulty in getting a result similar to (2.2). This is borne out by the following generalization of Lemma 2.1.

LEMMA 3.1. *For every ϵ, suppose there exists a convex polyhedral cone $\mathcal{C}_\epsilon$, generated by non-zero vectors $\beta_{\epsilon j}$, $(j = 1, \cdots, N(\epsilon))$, such that*

$$\mathcal{C}(\alpha_\epsilon) \subset \mathcal{C}_\epsilon \subset [\mathcal{C}(\pi_\epsilon)]^+. \tag{3.1}$$

Let $r_{\epsilon j}(t) = (\beta_{\epsilon j}, \pi_\epsilon(t))$, for every t. Then, for all ϵ and μ and for $j = 1, \cdots, N(\epsilon)$, $k = 1, \cdots, N(\mu)$, there exist constants $m_{\epsilon j,\mu k}$ such that, for all t,

$$r_{\epsilon j}(\mu t) = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} r_{\mu k}(t). \tag{3.2}$$

Further, the m's in (3.2) can be chosen to be non-negative if, and only if, the vector $\beta_{\epsilon j}A_{\epsilon\mu}$ belongs to $\mathcal{C}_\mu$.

PROOF. Lemma 1.1 shows that $\mathcal{C}_\mu$ has dimension $n(\mu)$. That is, there are $n(\mu)$ linearly independent $\beta_{\mu k}$'s. The linear span of the $\beta_{\mu k}$'s is therefore the whole of $E^{n(\mu)}$. The vector $\beta_{\epsilon j}A_{\epsilon\mu}$ lies in $E^{n(\mu)}$. Hence there are constants $m_{\epsilon j,\mu k}$ such that

$$\beta_{\epsilon j}A_{\epsilon\mu} = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k}\beta_{\mu k}. \tag{3.3}$$

(3.2) now follows by taking inner product with $\pi_\mu(t)$ and using (1.5) on the left side. It may also be noted that (3.2) implies (3.3). This is a consequence of the fact that $\mathcal{C}(\pi_\mu)$ has full dimension. The second assertion of the lemma is then immediate. This completes the proof of the lemma.

This lemma immediately yields a generalization of Theorem 2.1.

THEOREM 3.1. *Let $\{Y_n, n \geqq 1\}$ be a stationary process with a finite state-space. Let $\sum n(\epsilon) < \infty$ and suppose, for every ϵ, there exists a convex polyhedral cone $\mathcal{C}_\epsilon$ generated by non-zero vectors $\beta_{\epsilon j}$, $(j = 1, \cdots, N(\epsilon))$, such that $\mathcal{C}(\alpha_\epsilon) \subset \mathcal{C}_\epsilon \subset [\mathcal{C}(\pi_\epsilon)]^+$. For every ϵ, μ and for $j = 1, \cdots, N(\epsilon)$, assume that $\beta_{\epsilon j}A_{\epsilon\mu}$ belongs to $\mathcal{C}_\mu$. Then $\{Y_n\}$ is a function of a stationary Markov chain with $\sum N(\epsilon)$ states.*

PROOF. We will use the notation of the previous lemma. As in Lemma 2.1 we can choose the β's in such a way that $r_{\epsilon j}(\phi) = 1$. The non-negative constants $m_{\epsilon j,\mu k}$ yielded by the previous lemma then satisfy, as before, $\sum_{\mu=0}^{D-1}\sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} = 1$. We can therefore construct a transition matrix M as in (2.7). The existence of non-negative constants $m_{\epsilon j}^{(0)}$ satisfying (2.8) follows from the assumption that $\mathcal{C}(\alpha_\epsilon) \subset \mathcal{C}_\epsilon$. Repeating the steps (2.9) and (2.10), a stationary initial distribution $\mathbf{m}$ for M can then be constructed. But then Lemmas 2.2 and 2.3 hold word for word. This completes the proof of the theorem.

If each $\mathcal{C}(\pi_\epsilon)$ is polyhedral, then by taking $\mathcal{C}_\epsilon = [\mathcal{C}(\pi_\epsilon)]^+$ it is seen that the assumptions of Theorem 3.1 are satisfied. [See (2.4)]. The sufficient conditions given by 3.1 are therefore weaker than those given by Theorem 2.1. That the condition "$\beta_{\epsilon j}A_{\epsilon\mu}$ belongs to $\mathfrak{C}_\mu$" is a crucial one is shown by the following lemma.

LEMMA 3.2. *Suppose, for every* ϵ, *there is a convex polyhedral cone* $\mathfrak{C}_\epsilon$, *generated by vectors* $\beta_{\epsilon j}$, $(j = 1, \cdots, N(\epsilon))$, *such that for every* ϵ, μ *and for* $j = 1, \cdots, N(\epsilon)$, *the vector* $\beta_{\epsilon j}A_{\epsilon\mu}$ *belongs to* $\mathfrak{C}_\mu$. *Then, for every* ϵ, $\mathfrak{C}(\alpha_\epsilon) \subset \mathfrak{C}_\epsilon \subset [\mathfrak{C}(\pi_\epsilon)]^+$ *if, and only if,*

(i) *for every* ϵ, $\alpha_\epsilon(\phi)$ *belongs to* $\mathfrak{C}_\epsilon$, *and*

(ii) *for every* ϵ *and* j, $(\beta_{\epsilon j}, \pi_\epsilon(\phi)) \geqq 0$.

PROOF. The "only if" part is immediate. For the "if" part, let $r_{\epsilon j}(t) = (\beta_{\epsilon j}, \pi_\epsilon(t))$. The assumptions imply the existence of non-negative $m_{\epsilon j,\mu k}$ such that

$$r_{\epsilon j}(\mu t) = \sum_{k=1}^{N(\mu)} m_{\epsilon j,\mu k} r_{\mu k}(t). \tag{3.4}$$

(ii) implies that $r_{\epsilon j}(\phi) \geqq 0$. Then (3.4), with $t = \phi$, shows that $r_{\epsilon j}(\mu) \geqq 0$, for every μ. Proceeding by induction it follows that $r_{\epsilon j}(t) \geqq 0$ for every t. Thus $\mathfrak{C}_\epsilon \subset [\mathfrak{C}(\pi_\epsilon)]^+$.

(i) shows that there exist non-negative constants $q_{\epsilon j}(\phi)$ such that $\alpha_\epsilon(\phi) = \sum_{j=1}^{N(\epsilon)} q_{\epsilon j}(\phi)\beta_{\epsilon j}$. Define $q_{\epsilon j}(s)$, for all s, by induction as follows

$$q_{\mu k}(s\epsilon) = \sum_{j=1}^{N(\epsilon)} q_{\epsilon j}(s) m_{\epsilon j,\mu k}. \tag{3.5}$$

Then $q_{\epsilon j}(s) \geqq 0$, for all s and the assertion $\mathfrak{C}(\alpha_\epsilon) \subset \mathfrak{C}_\epsilon$ follows from the fact that

$$\alpha_\epsilon(s) = \sum_{j=1}^{N(\epsilon)} q_{\epsilon j}(s)\beta_{\epsilon j}. \tag{3.6}$$

The proof of (3.6) follows the same lines as that of Lemma (2.2) and is therefore omitted. The proof of the present lemma is now complete.

In terms of an underlying Markov chain $\{X_n\}$ from which the distribution of the given process $\{Y_n\}$ can be obtained through a function f, the quantity $q_{\epsilon j}(s)$ introduced in the proof of the preceding lemma has the following interpretation. If s has length n, then $q_{\epsilon j}(s) = P[(f(X_1), \cdots, f(X_n)) = s, X_{n+1} = \epsilon j]$. Equation (3.5) is then a direct consequence of the Markov character of $\{X_n\}$.

We will now show that the sufficient conditions given by Theorem 3.1 are also necessary when $\{Y_n\}$ is a *regular* function of a finite Markov chain. Let e_ϵ denote the $n(\epsilon) \times 1$ matrix all of whose elements are equal to 1. Part II of Lemma 3.1 of [1] then tells us that, for every ϵ, we can find a non-singular matrix $U_{\epsilon\epsilon}$ of order $n(\epsilon)$ such that: (a) the first row v_ϵ of $U_{\epsilon\epsilon}$ is non-negative, (b) $U_{\epsilon\epsilon}e_\epsilon = \pi'_\epsilon(\phi)$ and (c) $M_{\epsilon\mu} = U_{\epsilon\epsilon}^{-1}A_{\epsilon\mu}U_{\mu\mu}$ is non-negative.

Let $\beta_{\epsilon j}$, $(j = 1, \cdots n(\epsilon))$, be the row vectors of $U_{\epsilon\epsilon}^{-1}$ and let $\mathfrak{C}_\epsilon$ be the convex cone spanned by the $\beta_{\epsilon j}$'s. Since $M_{\epsilon\mu}$ is non-negative, the relation $U_{\epsilon\epsilon}^{-1}A_{\epsilon\mu} = M_{\epsilon\mu}U_{\mu\mu}^{-1}$ shows that $\beta_{\epsilon j}A_{\epsilon\mu}$ belongs to $\mathfrak{C}_\mu$. It was shown in Section 2 of [1] that it is always possible to take $s_{\epsilon 1}$ to be the empty sequence. Therefore $\alpha_\epsilon(\phi) =$

$(1, 0, \cdots, 0) = v_\epsilon U_{\epsilon\epsilon}^{-1}$. This shows that $\alpha_\epsilon(\phi)$ belongs to $\mathcal{C}_\epsilon$. Finally, $U_{\epsilon\epsilon}^{-1}\pi_\epsilon'(\phi) = e_\epsilon$ which is non-negative. That is $(\beta_{\epsilon j}, \pi_\epsilon(\phi)) \geqq 0$.

Lemma 3.2 therefore shows that $\mathcal{C}(\alpha_\epsilon) \subset \mathcal{C}_\epsilon \subset [\mathcal{C}(\pi_\epsilon)]^+$. The conditions of Theorem 3.1 are thus necessary in the regular case. We have not been able to extend this result to the general case. It is however believed that the final solution of the problem of characterizing functions of finite Markov chains lies in this direction.

4. A counterexample. Gilbert [3] has mentioned the possibility that the class of processes which are functions but not regular functions of finite Markov chains may be empty. In this section, we show by an example that the class is not empty.

The term "pseudo-Markov matrix" will denote a square matrix each of whose rows adds to 1. Lemma 3.1 of [1] shows how probabilities connected with a stationary process with $\sum n(\epsilon) < \infty$ can be computed from pseudo-Markov matrices through a functional approach. This procedure will be followed in this section.

Consider the pseudo-Markov matrix

$$M_1 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 1-\lambda_1 \\ 0 & -\lambda_2 & 0 & 0 & 1+\lambda_2 \\ 0 & 0 & \lambda_1 & 0 & 1-\lambda_1 \\ 0 & 0 & 0 & -\lambda_3 & 1+\lambda_3 \\ (1-\lambda_1)/2, & c(1+\lambda_2), & (1-\lambda_1)/2, & -c(1+\lambda_3), & \lambda_1 + c(\lambda_3-\lambda_2) \end{bmatrix}$$

with stationary initial distribution $(.25, c \times .5, .25, -c \times .5, .5)$. Let the 5 states be numbered 01, 02, 11, 12 and 21. Let g be a function defined by $g(\epsilon i) = \epsilon$, $(i = 1, 2; \epsilon = 0, 1)$, and $g(21) = 2$. Then g gives rise to a 3-state function process for which we can compute the pseudo-probabilities of various eventualities. We will get a proper process as soon as we choose the λ's and c in such a way that these pseudo-probabilities are non-negative.

It should be noted that: (i) since only one state of the underlying process goes into the state 2 of the function process, we must have $p(s2t) = p(s2)p(2t)/p(2)$; (ii) for the function process, the pseudo-probability is zero for any sequence which passes from 0 to 1 or from 1 to 0 without passing through 2. Hence it is enough to ensure that pseudo-probabilities of the form $p(0^n)$, $p(1^n)$, $p(2^n)$, $p(0^n2)$, $p(1^n2)$, $p(20^n2)$ and $p(21^n2)$ are non-negative, where ϵ^n denotes a sequence of $n\epsilon$'s in succession. In the following formulae ϵ is either 0 or 1.

$$p(\epsilon^n) = (.25)[\lambda_1^{n-1} + (-1)^{n-1+\epsilon} 2c\lambda_{(2+\epsilon)}^{n-1}],$$

$$p(2^n) = (.5)[\lambda_1 + c(\lambda_3 - \lambda_2)]^{n-1},$$

$$p(\epsilon^n 2) = (.25)[\lambda_1^{n-1}(1-\lambda_1) + (-1)^{n-1+\epsilon} 2c\lambda_{(2+\epsilon)}^{n-1}(1+\lambda_{(2+\epsilon)})],$$

$$p(2\epsilon^n 2) = (.25)[\lambda_1^{n-1}(1-\lambda_1)^2 + (-1)^{n-1+\epsilon} 2c\lambda_{(2+\epsilon)}^{n-1}(1+\lambda_{(2+\epsilon)})^2].$$

These expressions show that we will get a proper process if we choose the λ's and c in such a way that

$$(4.1)\quad \begin{array}{lll} 0 < \lambda_i < 1, (i = 1, 2, 3); & \lambda_1 > \lambda_i, (i = 2, 3); & 0 < c < .5; \\ \lambda_1 + c(\lambda_3 - \lambda_2) > 0; & (1 - \lambda_1)^k > 2c(1 + \lambda_i)^k, (k = 1, 2; i = 2, 3). & \end{array}$$

It is clear that the Conditions (4.1) can be satisfied. For example, take $\lambda_1 = .5$, $\lambda_2 = .4$, $\lambda_3 = .3$ and $c = .06$. We get

$$\lambda_1 + c(\lambda_3 - \lambda_2) = .494$$

$$(1 - \lambda_1) = .5, \quad 2c(1 + \lambda_2) = .168, \quad 2c(1 + \lambda_3) = .156$$

$$(1 - \lambda_1)^2 = .25, \quad 2c(1 + \lambda_2)^2 = .2352, \quad 2c(1 + \lambda_3)^2 = .2028.$$

Thus the Conditions (4.1) are satisfied.

Now for $i = 1, 2$, let A_i, u_i be real numbers with $u_1 \neq u_2$ and $A_i \neq 0$. Let $\theta_n = A_1u_1^{n-1} + A_2u_2^{n-1}$. Then $\theta_1\theta_3 - \theta_2^2 = A_1A_2(u_1 - u_2)^2 \neq 0$. In the expression for $p(\epsilon^n)$, observe that $\lambda_1 \neq \lambda_{2+\epsilon}$ and $c \neq 0$. Hence, for $\epsilon = 0, 1$, $p(\epsilon)p(\epsilon^3) - [p(\epsilon^2)]^2 \neq 0$. Thus for our 3-state process, $n(0) = n(1) = 2$ and $n(2) = 1$. That is, $n(\epsilon) \leqq 2$ for each ϵ. Corollary 2.1 then shows that this process can be expressed as the function g of a 5-state Markov chain.

Define the function h by $h(0) = h(1) = 0$ and $h(2) = 1$ and consider the 2-state process which is the function hg of the 5-state Markov chain obtained above. This author has considered this 2-state process in [1]. It is shown there that this new process has $\sum n(\epsilon) = 4$ and that it cannot be expressed as a function of a 4-state Markov chain—that is, it is not a regular function of a Markov chain.

We have thus constructed a stationary process which is a function but not a regular function of a finite Markov chain.

5. Acknowledgments. The author is grateful to Professor Edward W. Barankin under whose guidance this work was carried out. Thanks are due to Mr. Peter J. Bickel and Mr. Ashok P. Maitra for checking certain sections of the manuscript.

REFERENCES

[1] Dharmadhikari, S. W. (1963). Functions of finite Markov chains. *Ann. Math. Statist.* **34** 1022–1032.

[2] Fox, Martin. (1962). Conditions under which a given process is a function of a Markov chain (abstract). *Ann. Math. Statist.* **33** 1206.

[3] Gilbert, Edgar J. (1959). On the identifiability problem for functions of finite Markov chains. *Ann. Math. Statist.* **30** 688–697.